

To: Kristen Evangelista <Kristen.S.Evangelista@noaa.gov>
From: Jeff de La Beaujardière, NOAA Data Management Architect
Subject: Comments on Draft RFP for NextGen 4-D Weather Data Cube
Date: 2011-08-22

Dear Ms Evangelista,

Below please find my technical comments on the draft RFP for the NextGen 4-D Weather Data Cube. Note that I am submitting these in my role as NOAA Data Management Architect, but that these comments constitute my individual opinion and have not been circulated to NOAA's Environmental Data Management Committee or its Data Management Integration Team.

Thank you,
Jeff de La Beaujardière, PhD
NOAA NESDIS/OSD/TPIO

Doc.	Line or Ref.	Comment
RFP	76-77	Text states scope is "defined, already existing set of meteorological products" yet line 174 implies there will be new products "with improved time and spatial resolution," and line 193 refers to "legacy and proposed NOAA weather products."
RFP	89	The "desired net-centric format" mentioned here and elsewhere, sometimes as "net-ready formats" (line 105), is not named or specified, either in the referenced section C.4.3 Net-Centric, or in Attachment G - Functional Requirements. See also comment regarding Att. G, FR-2.2.12.
RFP	89-90	The text suggests that "Legacy data formats" will be discussed in C.4.14 Data Formats (actual number is C.4.12), which in turn refers to Attachment G as the place where legacy formats are enumerated. Specific legacy data formats are actually described in Att. D - Data Dictionary. This reference should be corrected. A brief summary of the legacy formats should be provided in the main body of the RFP.
RFP	196, 204	"Timely" is an adjective, not an adverb.
RFP	215	What does weather information "sharing" mean beyond the data dissemination described in lines 212-214?
RFP	378-381	OGC Web Services are mentioned in the Data Formats section, but are really services with multiple operations, not all of which are for data transfer. OGC services should be discussed in a separate section.
RFP	378-381	The OPeNDAP protocol, in wide use at NOAA but perhaps not at FAA, is not mentioned. Is it forbidden?
RFP	997-998	Text says "patches and upgrades will be defined and distributed by the Cube." This non-trivial functionality, which is not the same as data distribution, is never mentioned in G-Functional Requirements.
RFP	1010	Text says "there will be no direct access by the general public" to the data and functionality provided by the Cube. This is unfortunate. At minimum, provision should be made for the potential of separately hosting some of the most useful services (format conversion, coordinate transformation, subscription management, etc.) for purposes other than NOAA-to-FAA data transfer.

Doc.	Line or Ref.	Comment
RFP	1013-1015	Text says "Data transport for the Cube will be through secure telecommunications infrastructures with no data traversing the public Internet." This is unrealistic and unreasonable. First of all, the RFP is for the contractor to develop and operate Services, not to build a new infrastructure. Secondly, the cloud hosting called for in the RFP requires data to get in and out of the cloud, which typically occurs via the Internet. Third, MR-1.1.9 in Att. G calls for the use of protocols including HTTP and FTP. Finally, if the intent is to use existing dedicated, secure channels now in use by operational NOAA systems, then there is conflict with the statement in C.2.2 that "none of these [legacy] systems or their communications links will be affected by, or otherwise be integral to, this activity." Perhaps the RFP meant to say that data transport would occur across Virtual Private Networks established within the existing telecom infrastructure. If not, this section, and the Functional Requirements, must be clarified to indicate exactly which secure channels will be furnished by the government and which shall be built by the contractor.
RFP	1030	Text says data must be delivered the "the <u>most</u> expedient and efficient manner." The superlative "most" is hard to prove and continually evolves. Better to say "a manner that meets or exceeds the performance requirements PR-5" (for example).
RFP	1095	Regarding data format translations between provider and consumer, the RFP should suggest and approach for handling cases in which legacy formats do not include all the information or metadata required by the exchange format.
RFP	1105	PIREPs are mentioned but not defined.
RFP	1126	UDDI and ebXML are mentioned but not defined.
RFP	1138	ISO 19115 should be mandated as the primary metadata standard, with conversion to legacy FGDC CSDGM when needed.
RFP	1195	Text says NextGen Cube work "may require coordinate with various standards organizations." This implies that Contractor must be, or become, a member of the Open Geospatial Consortium (because a NOAA contractor cannot use NOAA's OGC membership). This should be mentioned, and permission given to include the membership cost in the proposal budget. Membership at the Technical level would be ideal. Some subcontractors may also need membership, perhaps only at the Associate level, to view or comment on non-public OGC documents.
RFP	1191	Text says "the Contractor may modify" Interface Description Document (Att. E) with "government" approval. Approval for substantial deviation from this very promising IDD, mandating as it does a variety of open standards of current interest at NOAA, <u>must not</u> be left at the sole discretion of NextGen project management! As an expensive and critical NOAA project that NWS hopes will "serve as the baseline for future system designs" (lines 102-103), it is essential that experts NOAA-wide be consulted regarding changes that may affect interoperability. A NextGen technical advisory group, or the NOAA Environmental Data Management Committee (EDMC) , may be appropriate forums.
RFP	1426-1427	The modular development approach is laudable, and suggests that there should be prioritization of the Functional Requirements which are at present undifferentiated in their importance and scheduling.
RFP	1482-1484	This text should perhaps be moved after line 1427 to reinforce Government's role in approving the modular development schedule.
FR	1.1.3	This is two requirements--split them.
FR	1.2.2	Define "metadata tagging."
FR	1.2.3	Semantic translation should be a separate requirement from FR-1.2 Register Weather Metadata.
FR	1.2.4	Define "synchronize weather metadata."
FR	1.3.1	Explain "design-time discovery."
FR	1.3.5	Is this matching to be exact or inexact, in light of FR 1.3.3 & 1.3.4?

Doc.	Line or Ref.	Comment
FR	2	Requirements under FR-2 "Cube shall provide access to data and services" are poorly stated and mix concepts of data access, transformation between coordinate systems and units of measure, format conversion, caching, subsetting, and support for legacy data providers. Recommend major rewrite to separately discuss each of these topics as Level 1 requirements, with appropriate Level 2 & 3 requirements. I began such a revision but did not have time to finish before the comment deadline, and would be glad to discuss further.
FR	2.1.4	"NOAA-internal" implies a separate security requirement.
FR	2.2.21	FR-2.2.21 is redundant with 2.2.7. Also, EPSG is not a CRS. Rather, European Petroleum Survey Group maintains a database of CRS identifiers and parameters. The EPSG DB includes many obscure CRS; NextGen probably does not wish to mandate arbitrary conversion between any two.
FR	CR-2	Cube services are not intended to "contain" data, but rather to contain metadata and to enable access to data stored elsewhere.
FR	PR-3	This is not a requirement, but rather a contractual statement about requirements.
FR	PR-6.2	2 hours is over-long for adjusting to increased demand in a cloud-deployment scenario. Amazon EC2, for example, claims minutes rather than hours for scaling up or down.
FR	2.2.12	Text mentions some but is vague and somewhat incorrect. "Climate & Forecast formats" are mentioned, which is presumably a reference to the Climate & Forecast Conventions that are applicable to files in NetCDF format (http://cf-pcmdi.llnl.gov/). NetCDF4 is mentioned; it is not clear whether this forbids NetCDF3 (in wider use at this time). The phrasing of the text states there will be conversion between specific pairs of formats (e.g., NetCDF4 and GRIB2) and implies that other conversions to or from those formats is not needed.
IDD	All	Contractor should be encouraged to consider the NSF Ocean Observatories Initiative (OOI) Cyberinfrastructure design, which draws heavily on concepts of subscription, cloud hosting for scalability, and service adaptors.
IDD	R34	WSDL is mandatory, but not all OGC Web Services have WSDL descriptions.
IDD	R38	Use of OGC WFS for non-gridded observations (points, profiles, moving sensor) is mandated. OGC Sensor Observation Service (SOS) should be considered for this purpose.
IDD	R72, R75, R80	Publication/subscription extensions for WFS, WCS WMS are called for. Instead, should consider a separate pub/sub layer in the architecture that provides subscription management functionality for core services that, as presently defined, operate only in "pull" mode.
IDD	2.3.4.2.4	Unless a NOAA customer needs it, JMBL should be deferred until FAA support requires it.